

Name: Revanth Penugonda

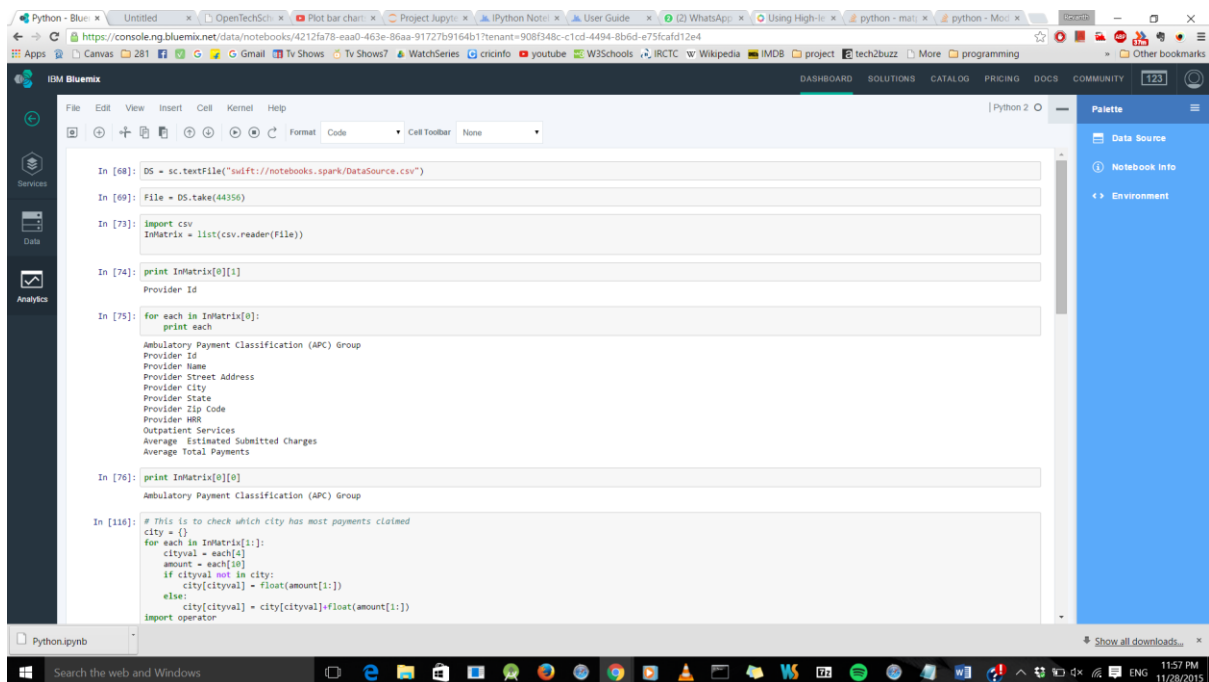
ID: 010690095

Extra Credit Assignment

Github Link: https://github.com/revaries/CMPEC72_ExtraCredit.git

Steps to Reproduce:

- Create an Apache Spark Service in IBM Bluemix
- Create a new Instance of Spark and also the Object Data Store
- Create a New notebook with python
- Drag and drop the data object



The screenshot shows a Jupyter Notebook running in the IBM Bluemix console. The notebook contains the following Python code:

```
In [68]: DS = sc.textFile("s3://notebooks-spark/DataSource.csv")
In [69]: File = DS.take(44356)
In [73]: import csv
In [73]: InMatrix = list(csv.reader(File))
In [74]: print InMatrix[0][1]
Provider Id
In [75]: for each in InMatrix[0]:
    print each
Ambulatory Payment Classification (APC) Group
Provider Id
Provider Name
Provider Street Address
Provider City
Provider State
Provider Zip Code
Provider HRR
Outpatient Services
Average Estimated Submitted Charges
Average Total Payments
In [76]: print InMatrix[0][0]
Ambulatory Payment Classification (APC) Group
In [116]: # This is to check which city has most payments claimed
city = {}
for each in InMatrix[1:]:
    cityval = each[4]
    amount = each[10]
    if cityval not in city:
        city[cityval] = float(amount[1:1])
    else:
        city[cityval] = city[cityval]+float(amount[1:1])
import operator
```

The interface includes a top navigation bar with 'Python - Blue' and 'Untitled' tabs, a left sidebar with 'Services', 'Data', and 'Analytics' sections, and a right sidebar with 'Data Source', 'Notebook Info', and 'Environment' sections. The bottom status bar shows the time as 11:57 PM on 11/28/2015.

[illegible]